



AIR

清华大学智能产业研究院

Institute for AI Industry Research, Tsinghua University

AI 大模型时代

Large Models, ChatGPT, Generative AI, and Education

张亚勤

清华大学讲席教授、智能产业研究院 (AIR) 院长

2023年6月

目录



AIR

清华大学智能产业研究院

Institute for AI Industry Research, Tsinghua University

1 AI大趋势

2 ChatGPT现象

3 对教育的思考



AIR

清华大学智能产业研究院
Institute for AI Industry Research, Tsinghua University

数字化 3.0: 信息+物理+生物世界数字化



基建: 道路、城市...

交通: 车辆、飞机...

产业: 工厂、机床...

生活: 家居、生活服务...



大脑、身体、器官...

生物电子芯片、脑机接口 ...

基因、蛋白质...

细胞、分子...

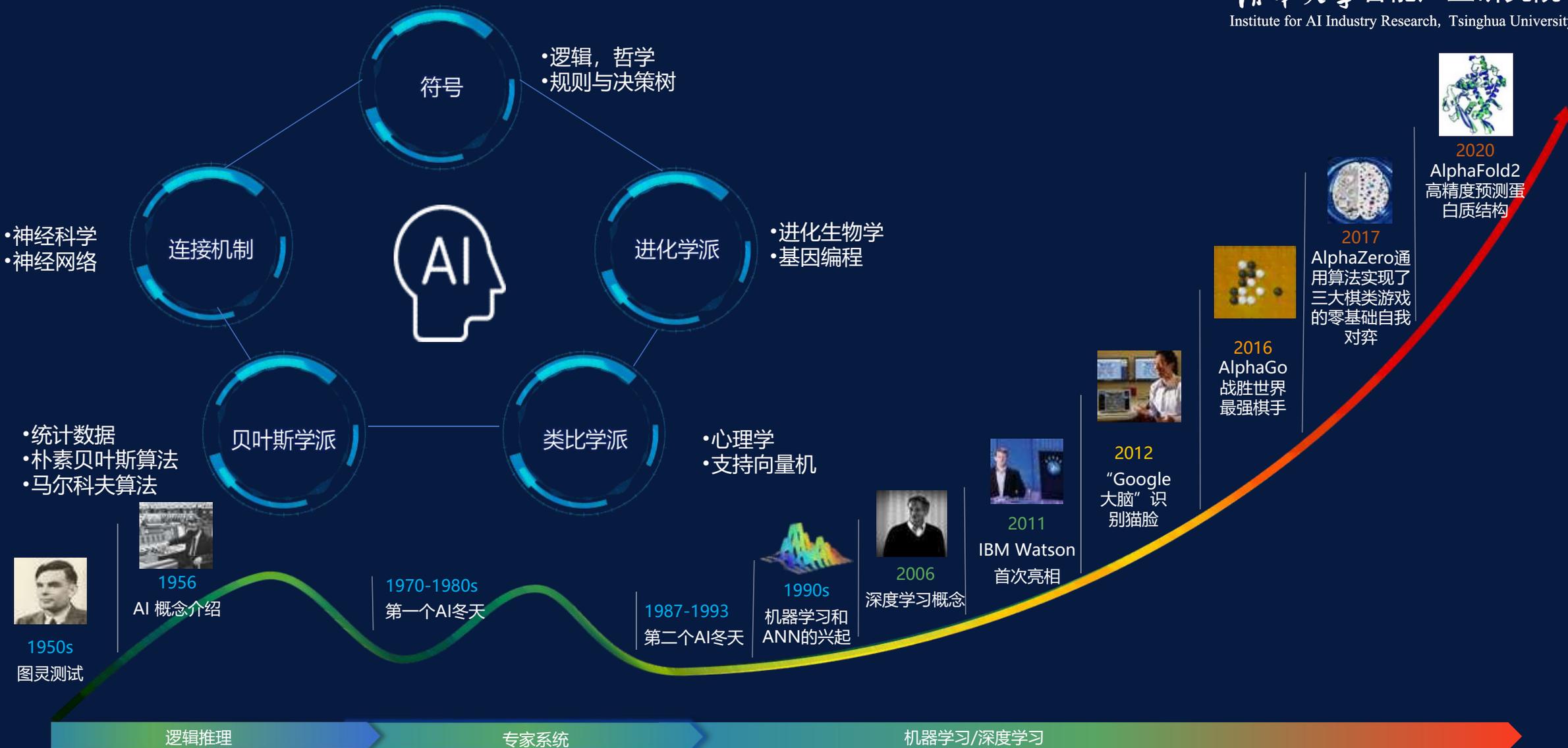
数字化1.0 & 2.0: 原子→比特

数字化3.0: 比特 ↔ 原子+分子

人工智能：机器学习智能和知识的探索



清华大学智能产业研究院
Institute for AI Industry Research, Tsinghua University



(The Master Algorithm, Pedro Domingos, 2017)

AI/深度学习 → 预训练, 多模态, 大模型, 生成式

透明性, 可解释, 因果性, 安全, 隐私, 伦理



清华大学智能产业研究院
Institute for AI Industry Research, Tsinghua University



生成式AI高速发展



清华大学智能产业研究院
Institute for AI Industry Research, Tsinghua University

- “AI gets creative” 入围 *Science* 2022年年度科学突破

太空歌剧院

由艺术家Jason Allen使用文本到图像生成AI平台Midjourney制作，获得美国科罗拉多州博览会艺术比赛数字类别头奖



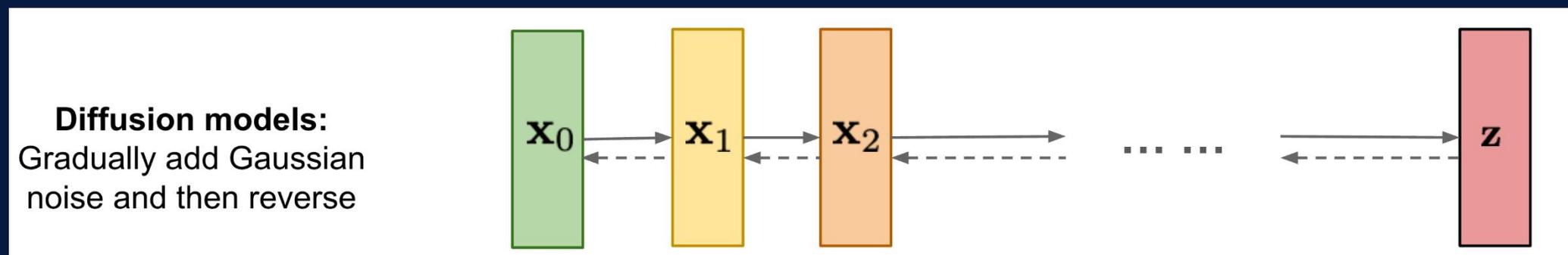
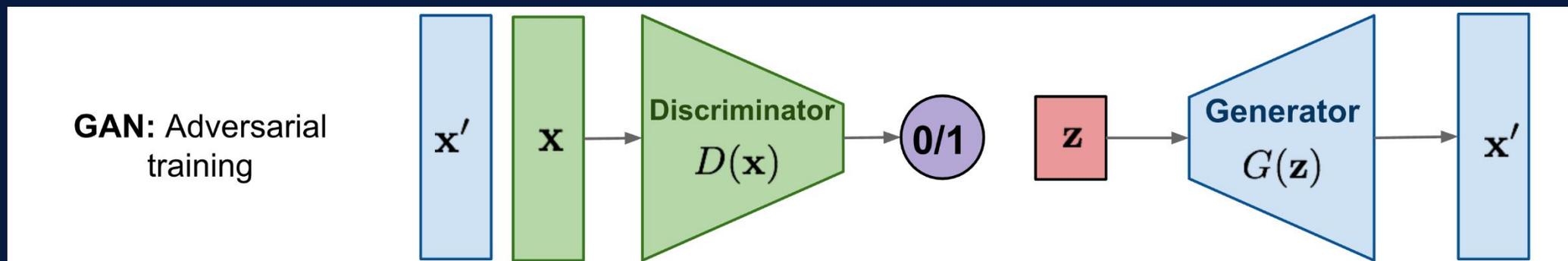
(来源: https://www.science.org/content/article/breakthrough-2022#section_ai)

AIGC: 图像生成



清华大学智能产业研究院
Institute for AI Industry Research, Tsinghua University

- GAN和Diffusion Models是最流行的图像生成模型
- Diffusion Models逐渐占主导地位，产生Midjourney, Runway, Stable Diffusion等



(图片来源: <https://lilianweng.github.io/posts/2021-07-11-diffusion-models/>)

AIGC: 由文本描述生成图像



清华大学智能产业研究院
Institute for AI Industry Research, Tsinghua University

- 代表性平台Midjourney生成的图像真假难辨



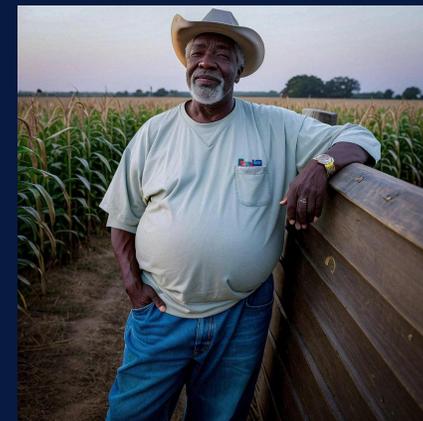
(图片来源: <http://yesaiwen.com/midjourney-images-with-prompts>)

AIGC: 可控图像生成



清华大学智能产业研究院
Institute for AI Industry Research, Tsinghua University

- ControlNet根据指定人物姿态，生成性别、衣着、环境各异的逼真图像



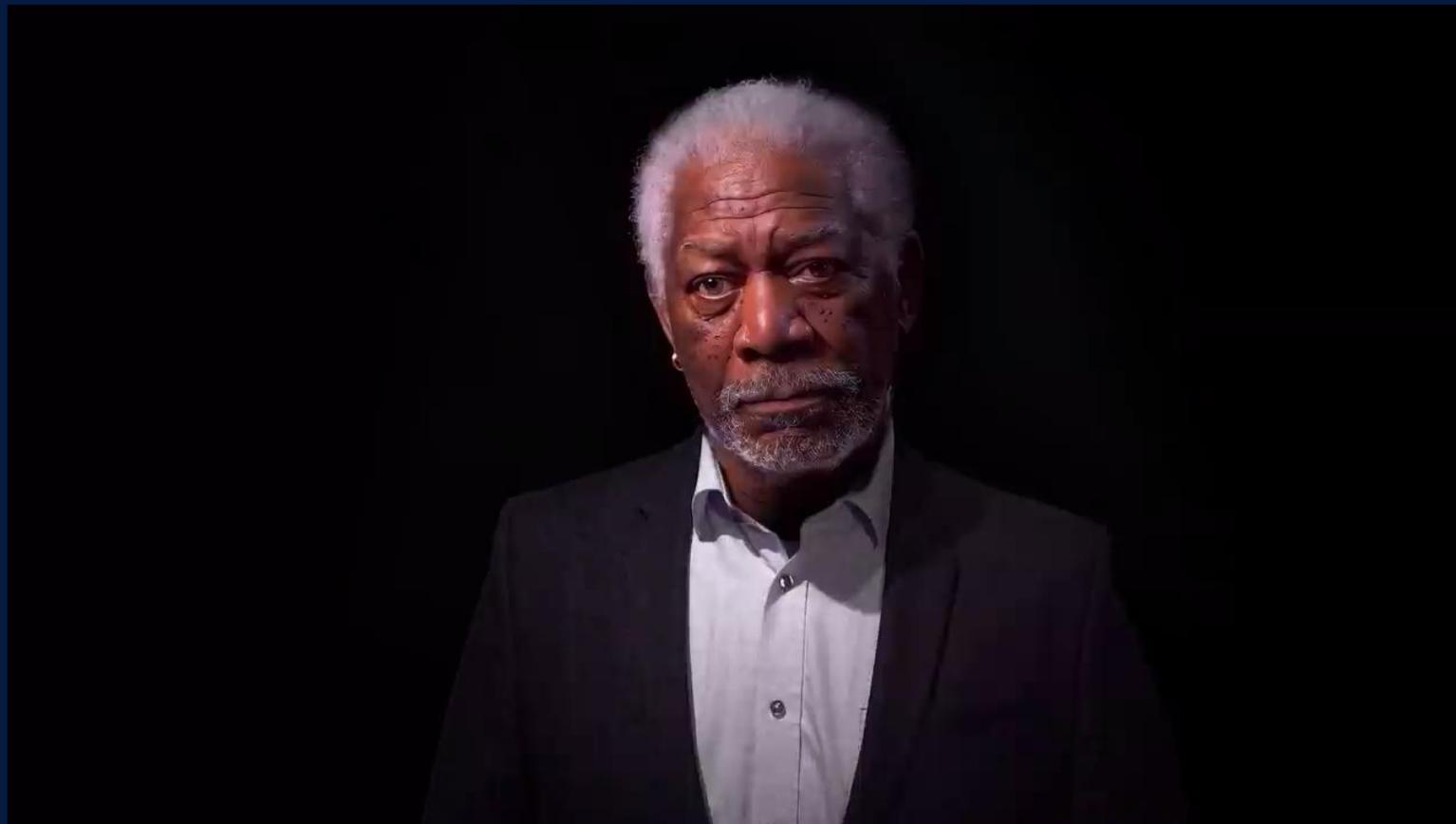
(图片来源: https://www.reddit.com/r/StableDiffusion/comments/1142dt/controlnet_magicposer_app_realistic_vision)

AIGC: DeepFake



清华大学智能产业研究院
Institute for AI Industry Research, Tsinghua University

- 合成高度逼真的人像或视频



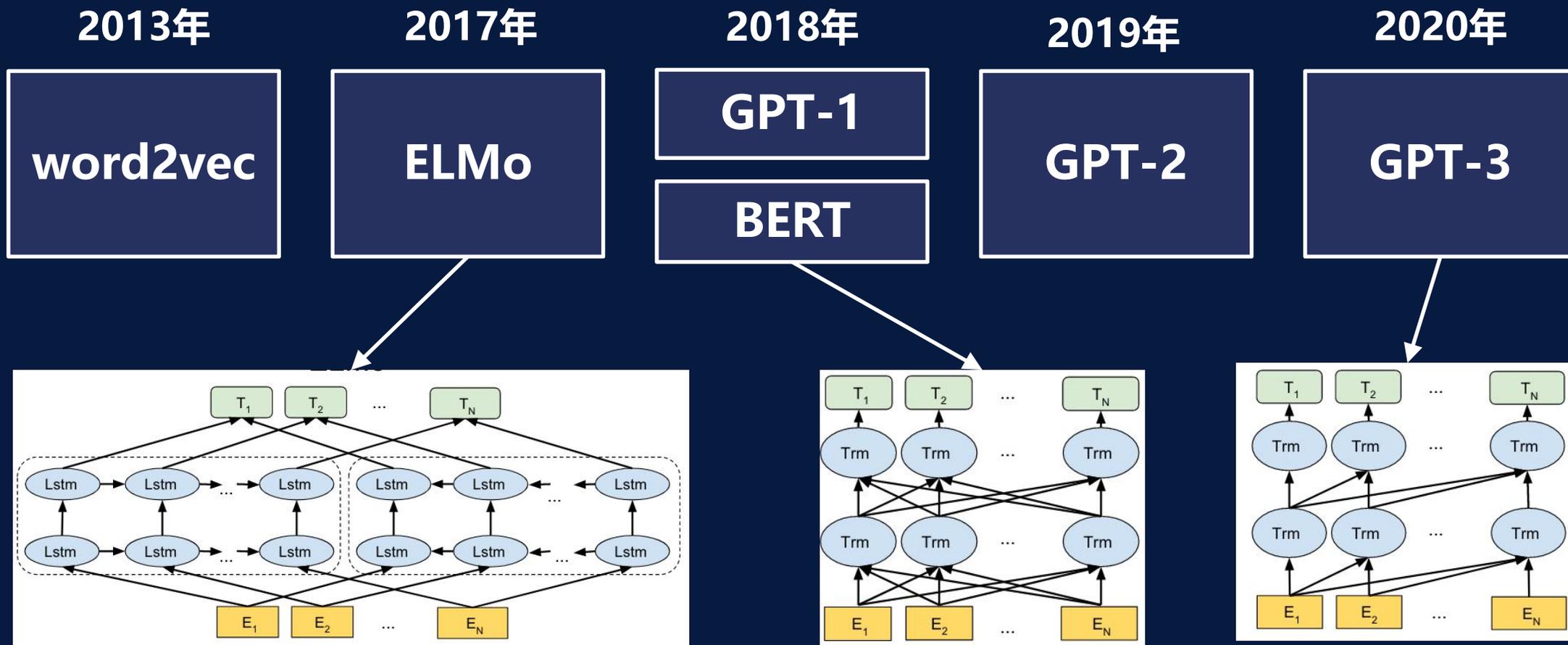
(视频来源: https://www.youtube.com/watch?v=oxXpB9pSETo&ab_channel=DiepNep)

AIGC: 语言大模型



清华大学智能产业研究院
Institute for AI Industry Research, Tsinghua University

- 从词向量预训练转向双向语言模型，并最终转向生成式预训练



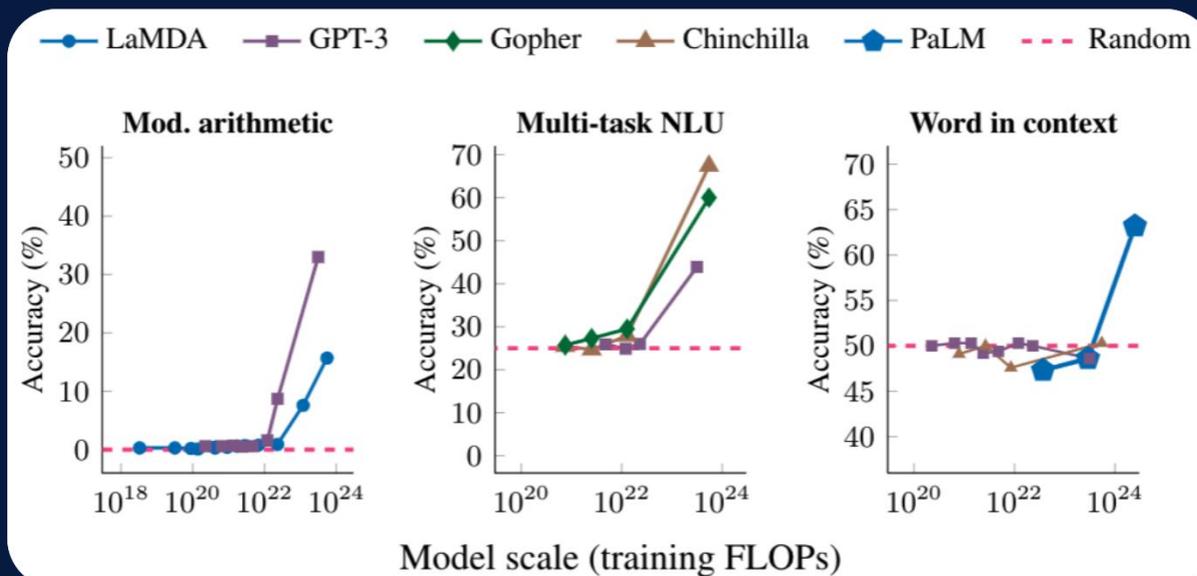
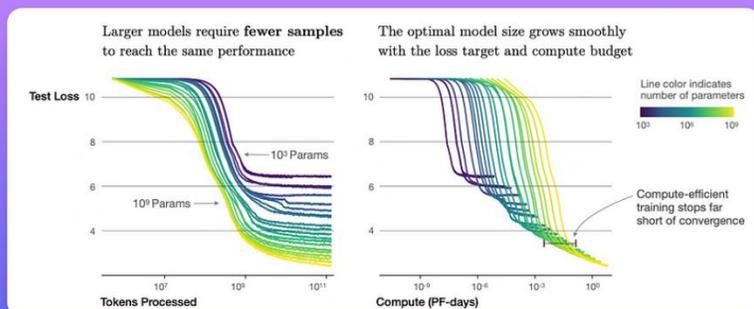
(Devlin et al., 2018)

基础大模型 (ChatGPT+, AIGC)

- 预训练, 自监督学习 (MASK), 强生成能力, 自我进化 (GPT 4 ==》 AutoGPT, GPT4+, ...))
- 泛化能力 (PLLM+SFT → 指令微调 + 多项下游任务)
- 人类反馈强化学习 (RLHF) -- 训练奖励函数 ==》 对齐
- 涌现和统一 (Emergence and Homogenization) : 人的大脑工作原理

规模效应 带来质变, Token-based Transformer

Scaling Laws for Neural Language Models



Model scale (training FLOPs)

ChatGPT现象



清华大学智能产业研究院
Institute for AI Industry Research, Tsinghua University

- **人类历史上最受欢迎的产品：2个月1亿用户（2022年12-2023年1月底）**
- **全球范围影响力巨大：远超AlphaGo**
- **提供用户界面：不完美但高价值**
- **获得各界高度评价，诞生若干“时刻”**
 - **Anderson：“网景”时刻**
 - **Huang：“iPhone”时刻**
 - **Gates：“GUI时刻”，地位类似PC、互联网**
 - **.....**

我的“Chat GPT”时刻！

• 通过“图灵测试”

- 对话式AI (conversational AI) / 聊天机器人已有50+年研究历史
- 第一个智能体 (generative AI) 通过“图灵测试”

```
Welcome to
          EEEEEEE LL      IIII  ZZZZZZ  AAAAA
          EE       LL      II     ZZ     AA  AA
          EEEEEEE LL      II     ZZZ     AAAAAAA
          EE       LL      II     ZZ     AA  AA
          EEEEEEE LLLLLL  IIII  ZZZZZZ  AA  AA

Eliza is a mock Rogerian psychotherapist.
The original program was described by Joseph Weizenbaum in 1966.
This implementation by Norbert Landsteiner 2005.

ELIZA: Is something troubling you ?
YOU:   Men are all alike.
ELIZA: What is the connection, do you suppose ?
YOU:   They're always bugging us about something or other.
ELIZA: Can you think of a specific example ?
YOU:   Well, my boyfriend made me come here.
ELIZA: Is it important to you that your boyfriend made you come here ?
YOU:   He says I'm depressed much of the time.
ELIZA: I am sorry to hear that you are depressed.
YOU:   It's true. I am unhappy.
ELIZA: Can you explain what made you unhappy ?
YOU:
```

Siri

Cortana

Alexa

Google

小度

天猫精灵

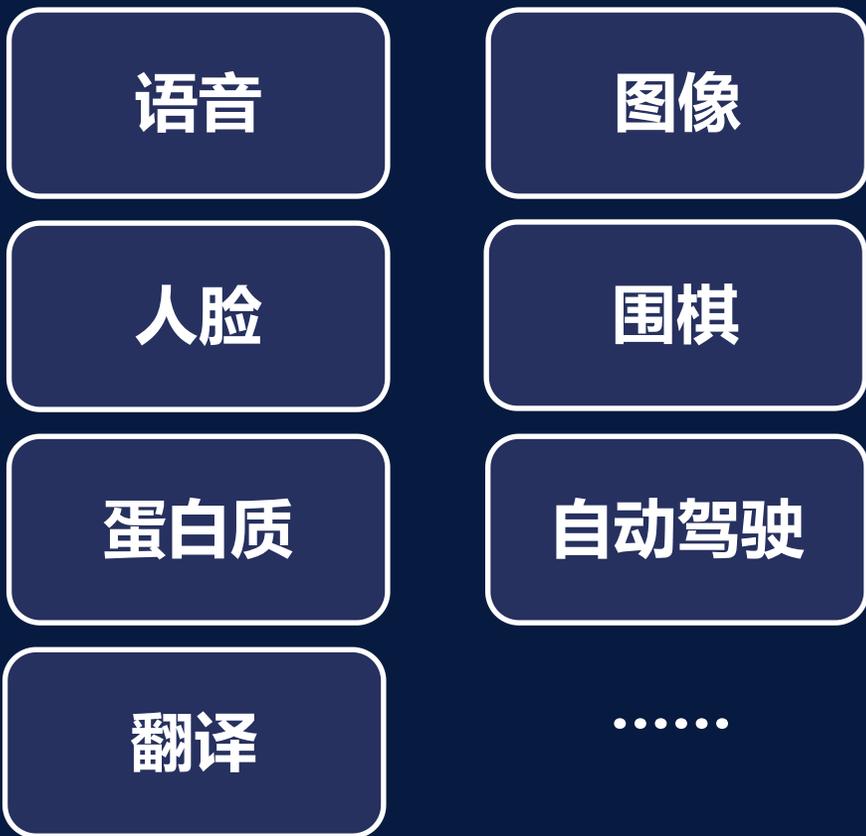
ELIZA (伊莉莎) --Joseph Weizenbaum
MIT 1966年，临床治疗中模仿心理医生



我的“Chat GPT”时刻！

- 开启了通用人工智能之路（深度学习2.0）

专用任务



通用任务





我的“Chat GPT”时刻！

- GPT+：人工智能时代的操作系统 OS → 重构/重塑

PC互联时代	Intel x86 + Windows + 浏览器 + 应用	1倍
移动互联时代	ARM + iOS/Andriod + 应用商店 + APP	10倍
智能时代	GPU/CPU/XPU + 基础大模型 + IAAS + 垂直模型 + SAAS	100倍

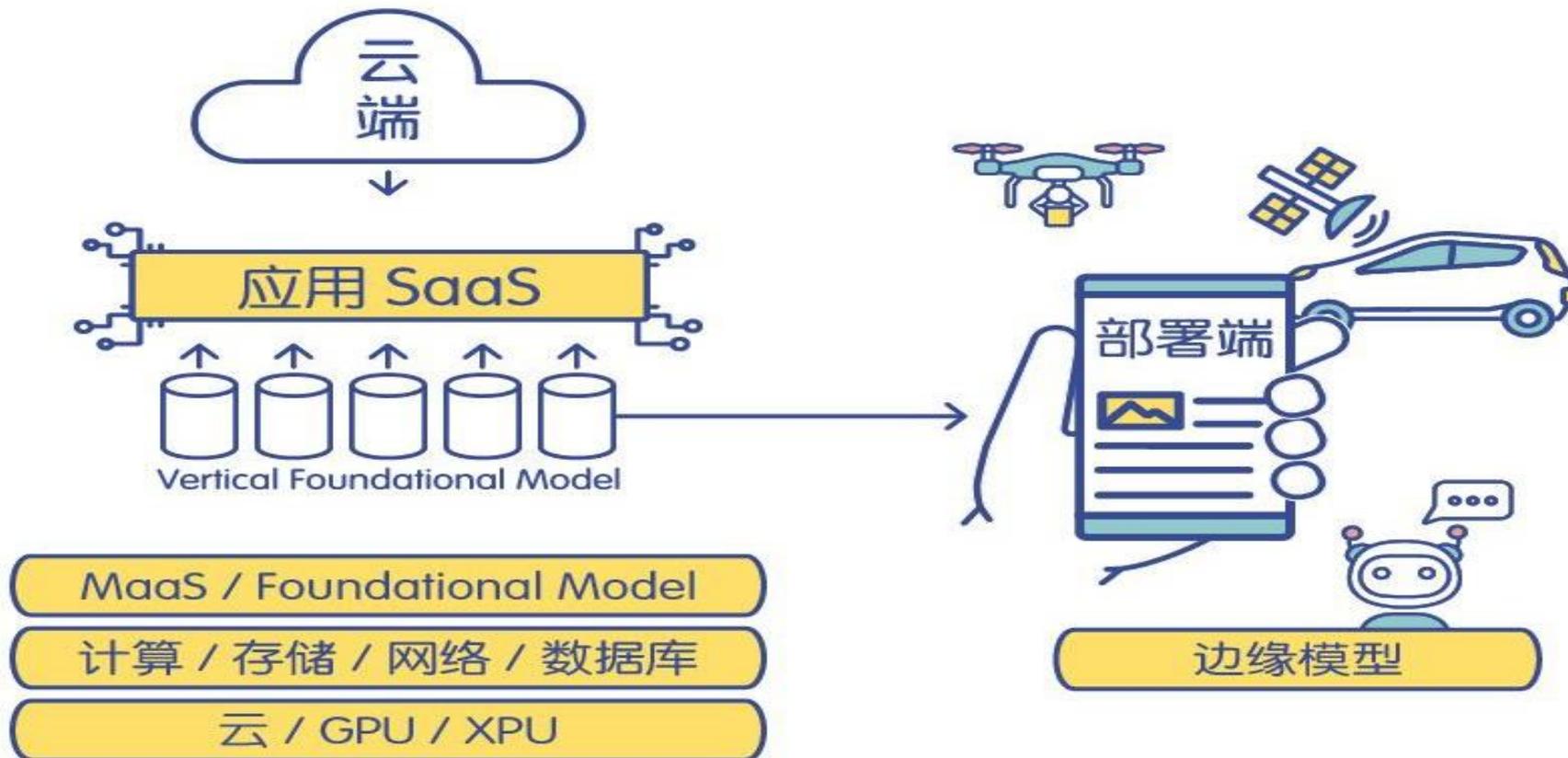
未来的技术架构



AIR

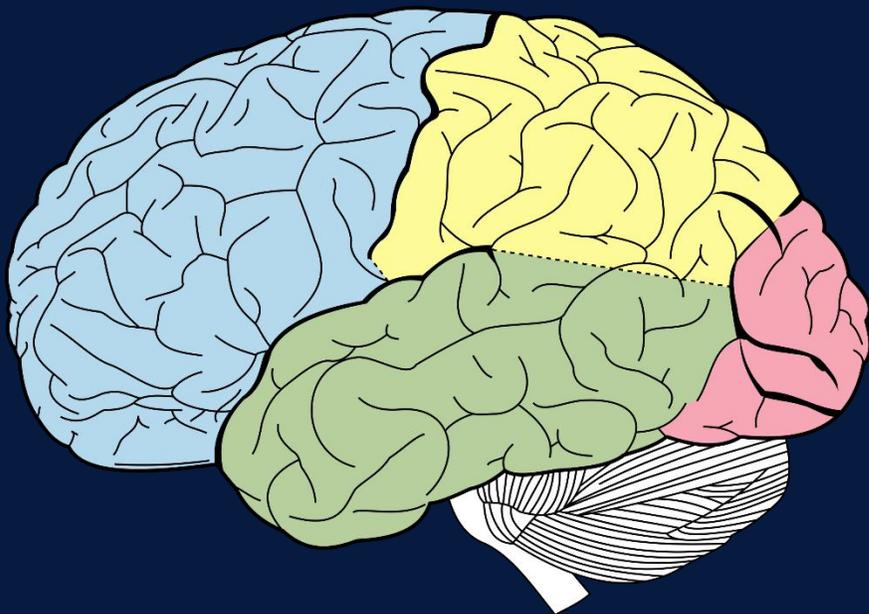
清华大学智能产业研究院

Institute for AI Industry Research, Tsinghua University



未来方向：效率

- 当前的大模型稠密激活，计算效率远低于人脑，商用成本高昂
- 改进方向：稀疏激活网络、效果更优的小网络等等



1.2 ~ 1.6kg, 10^{15} 突触连接, 约20瓦

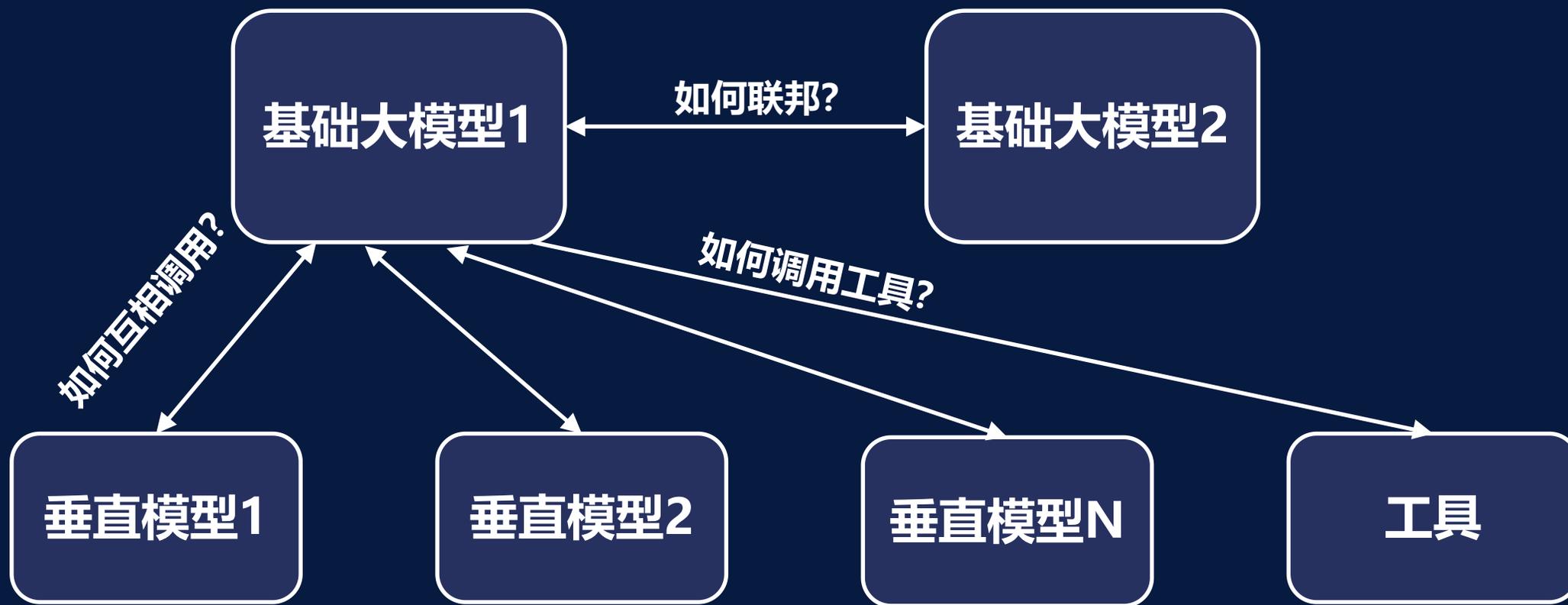


10000块V100训练14.8天, 单块V100 250瓦



未来方向：模型交互

- 模型间如何互相调用？模型如何使用工具？
- 基础模型间如何进行联邦学习？





未来方向：边缘部署

- 边缘端算力有限，如何在边缘端利用基础大模型？



如何高效推理、学习？

AIGC大模型将带来新范式——生成AI



清华大学智能产业研究院
Institute for AI Industry Research, Tsinghua University

Discriminative AI → Generative AI:

Generate
Image,
Video,
Speech,
Music

Generate
Text
Code
Tools

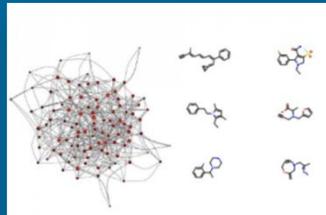
Generate
Molecule
Drugs
Materials
Equations
(AI4S)



Image generation
(by StyleGAN, From Internet)



Text generated by GPT-2,
Examples from Internet



Graph Neural Networks

What ?

How ?

Why?

通用人工智能



总结： AI大模型时代

- 多算法/多任务/多模型 → 统一算法/多模态/预训练大模型
- ChatGPT: 通过“图灵测试”，通识，开启通用人工智能时代亮光
- 基础大模型：Token-based, Scaling Law, 涌现和统一，AI时代的操作系统
- 行业垂直基础模型+开源模型 + API/Plugin /Auto → 新2C/2B应用新生态（100X）
- 创新和教育的巨大机遇！

A few more Words to Students



清华大学智能产业研究院
Institute for AI Industry Research, Tsinghua University

- 做永恒的学习者，应对变化
- 独特的观点，批判性思维
- 大胆试错，找到你的路径
- 保持激情、纯真、棱角和锋芒



AIR

清华大学智能产业研究院

Institute for AI Industry Research, Tsinghua University

**When we dare to dream,
we achieve greatness.**



扫一扫 关注我们